AD

Technical Memorandum 16-89

PROXIMITY COMPATIBILITY AND INFORMATION DISPLAY: THE EFFECTS OF SPACE AND COLOR ON THE ANALYSIS OF AIRCRAFT STALL CONDITIONS

Anthony D. Andre
Christopher D. Wickens

Aviation Research Laboratory and Department of Psychology
University of Illinois

October 1989
AMCMS Code 611102.74A0011

Approved for public release;
distribution is unlimited.

U.S. ARMY HUMAN ENGINEERING LABORATORY
Aberdeen Proving Ground, Maryland

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE

REPORT DOCUMENTATION PAGE

Form Approved OMB No. 0704-0188

1a. REPORT SECURITY CLASSIFICATION	1b. RESTRICTIVE MARKINGS
2a. SECURITY CLASSIFICATION AUTHORITY Unclassified	3. DISTRIBUTION/AVAILABILITY OF REPORT Approved for public release; distribution is unlimited.
2b. DECLASSIFICATION/DOWNGRADING SCHEDULE	
4. PERFORMING ORGANIZATION REPORT NUMBER(S)	5. MONITORING ORGANIZATION REPORT NUMBER(S) Technical Memorandum 16-89
6a. NAME OF PERFORMING ORGANIZATION University of Illinois at Urbana - Champaign 6b. OFFICE SYMBOL (If applicable)	7a. NAME OF MONITORING ORGANIZATION Human Engineering Laboratory
6c. ADDRESS (City, State, and ZIP Code) 1 Airport Road Savoy, Illinois 61874	7b. ADDRESS (City, State, and ZIP Code) Aberdeen Proving Ground Maryland 21005-5001
8a. NAME OF FUNDING/SPONSORING ORGANIZATION 8b. OFFICE SYMBOL (If applicable)	9. PROCUREMENT INSTRUMENT IDENTIFICATION NUMBER
8c. ADDRESS (City, State, and ZIP Code)	10. SOURCE OF FUNDING NUMBERS

PROGRAM ELEMENT NO.	PROJECT NO.	TASK NO	WORK UNIT ACCESSION NO
6.11.02	1L161102B74	A	

11. TITLE (Include Security Classification)
Proximity Compatibility and Information Display: The Effects of Space and Color on the Analysis of Aircraft Stall Conditions

12. PERSONAL AUTHOR(S)
Andre, Anthony D.; Wickens, Christopher D. -- University of Illinois at Urbana - Champaign

13a. TYPE OF REPORT	13b. TIME COVERED	14. DATE OF REPORT (Year, Month, Day)	15. PAGE COUNT
Final	FROM ____ TO ____	1989, October	43

16. SUPPLEMENTARY NOTATION

17. COSATI CODES

FIELD	GROUP	SUB-GROUP
01	03	
05	08	

18. SUBJECT TERMS (Continue on reverse if necessary and identify by block number)
display; color; clutter; attention; aviation; spacial organization; information processing; compatibility

19. ABSTRACT (Continue on reverse if necessary and identify by block number)

The proximity compatibility principle (Wickens, 1987) asserts that when a task requires the integration of multiple sources of information, performance will be best supported when that information is displayed in close proximity. Conversely, when a task requires attention to be focused on a specific source of information, performance will be best supported by a more separated display. To assess the validity of this principle, a series of three experiments were conducted in which subjects monitored a display of flight parameters critical to aircraft stability and were required to either predict the likelihood of an aircraft stall (information integration) or to recall the value of a single flight parameter (focused attention). Display proximity of relevant information was imposed through spatial closeness and color similarity. The results indicate that color adheres to the proximity compatibility principle, but that space does not. Instead, the spatial proximity between relevant and irrelevant information appears to be the dominant factor affecting performance across both tasks. The data are discussed in terms of their practical implications for multi-element display design.

20. DISTRIBUTION/AVAILABILITY OF ABSTRACT ☐ UNCLASSIFIED/UNLIMITED ☒ SAME AS RPT ☐ DTIC USERS	21. ABSTRACT SECURITY CLASSIFICATION Unclassified
22a. NAME OF RESPONSIBLE INDIVIDUAL Technical Reports Office	22b. TELEPHONE (Include Area Code) (301) 278-4478 — 22c. OFFICE SYMBOL SLCHE-SS-IR

DD Form 1473, JUN 86 — Previous editions are obsolete

SECURITY CLASSIFICATION OF THIS PAGE

UNCLASSIFIED

AMCMS Code 611102.74A0011 Technical Memorandum 16-89

PROXIMITY COMPATIBILITY AND INFORMATION DISPLAY: THE EFFECTS OF SPACE AND COLOR ON THE ANALYSIS OF AIRCRAFT STALL CONDITIONS

Anthony D. Andre
Christopher D. Wickens

Aviation Research Laboratory and Department of Psychology
University of Illinois

October 1989

APPROVED: John D. Weisz
JOHN D. WEISZ
Director
Human Engineering Laboratory

Approved for public release;
distribution is unlimited.

U.S. ARMY HUMAN ENGINEERING LABORATORY
Aberdeen Proving Ground, Maryland 21005-5001

ACKNOWLEDGMENTS

This research was supported by a grant from the U.S. Army Human Engineering Laboratory (DAAA15-86-K-0013). Mr. Frank Malkin and Ms. Kathleen Christ were the technical monitors. The authors would like to thank Mr. Frank Miller for programming the displays used in these experiments.

CONTENTS

TABLE

PROXIMITY COMPATIBILITY AND INFORMATION DISPLAY: THE EFFECTS OF SPACE AND COLOR ON THE ANALYSIS OF AIRCRAFT STALL CONDITIONS

INTRODUCTION

Modern advances in science and technology have had a major impact on the human-machine interfaces of complex systems such as those found in the aircraft cockpit or process control monitoring station. In particular, the rapid development of computer technology has led to the widespread use of computer-based control, decision, and display systems that provide potentially safer and more efficient operation than previous or existing electro-mechanical systems.

Ironically, though, these advances in technology often contribute to the increased complexity of the systems they are intended to support (Rasmussen, 1986; Wiener & Curry, 1980). Furthermore, as the complexity of a system increases so too might the cognitive demands of the system operator whose role shifts away from active manual control and toward more mediational planning, monitoring, and communication activities (Curry, 1985; Huntoon, 1985; Moray, 1988; Wiener, 1985; Wierwille, Rahimi, & Casali, 1985). Accordingly, these aftereffects of new technology have influenced many facets of human-machine interface design in modern complex systems (Rasmussen, 1986). However, because they are often "...considered separately, by different persons and at different phases of system design" (Rasmussen, 1986), the design community is becoming increasingly aware of the need for an integrated approach to system design. To accomplish this, human engineering research must assess the extent to which several factors interact to influence human performance in complex systems. Nowhere is this assessment more important than in the cockpit of a high-performance aircraft, where the benefits of new technology will not be realized unless the associated displays are designed to be compatible with the information requirements of the operator in control (Lovesey, 1986).

Over the past 20 years, the information requirements imposed on the pilots of high-performance aircraft have been continually increasing (Statler, 1984; Stokes & Wickens, 1988) while the available space to provide this information has either remained the same, or in some cases, decreased. The pilot is now faced with the increasingly difficult task of extracting, integrating, and acting upon critical information within a short period of time (Kramer, Wickens, Goettl, & Harwood, 1986). As a result, human engineering research has sought new solutions for supporting the complex information processing typically required in aircraft cockpits and other large-scale human-machine interfaces. The main component of this effort has been the consideration of the nature of the operator's cognitive task requirements. In turn, these requirements provide the basis for an information processing approach whereby the capabilities and limitations of human perception and information processing guide the development of displays to support complex task performance.

Recent theoretical developments in the display-cognitive interface have provided useful guidance for improving human information extraction and processing performance with multi-element display interfaces. In particular, the proximity compatibility principle (Boles & Wickens, 1987; Carswell & Wickens, 1987; Wickens, 1987) has been proposed as a means of guiding the development of display formats that support human information processing requirements. This principle asserts that the degree to which multiple channels of information are similarly displayed (i.e., displayed in proximity) should directly correspond to the degree to which the appropriate task requires similar or integrated processing of these information channels (Carswell & Wickens, 1988).

This report describes a series of three experiments that are part of a larger research effort to validate the proximity compatibility principle as a useful guideline for the design of complex, multicue information environments such as the aircraft cockpit. The purpose of this report is (1) to highlight the need for an information processing approach to principle-driven display design, and (2) to assess the extent to which two representations of display proximity (closeness in space and similarity in color) adhere to the proximity compatibility principle.

AN INFORMATION PROCESSING APPROACH TO DISPLAY DESIGN

Visual Displays

Most of the information in a cockpit is conveyed through visual displays. System designers must therefore ensure that these displays are easy to locate and easy to read. Consequently, the visual considerations of the pilot's task are at the center of attention in cockpit designs (Lyons & Roe, 1980). With the advent of the computer, and more specifically the electronic displays currently used in several aircraft, system designers now possess unheralded flexibility in the manner they can display information to the pilot. These developments have provided the designer with many design options previously unavailable (Laycock & Chorley, 1980). Some of these options include the ability to manipulate the location, color, size, and shape of the displayed information. Accordingly, this new technology has become a major component in the formatting of multi-element electronic displays.

Displays that are "technology driven," however, do not always facilitate improved levels of processing and performance (Huntoon, 1985; Wickens, 1988). Color coding with many advanced cathode-ray tube (CRT) displays, for example, can be used to greatly enhance the aesthetic appearance of the display attributes without effectively supporting performance (Banbury, 1984; Wickens, 1988). Although the perceptual characteristics of display formats (e.g., position, shape, color, orientation) play an important role in the perception and subsequent processing of the display attributes (Cleveland, 1985; Cleveland & McGill, 1984), the nature of the cognitive and response operations carried out on the perceived variables plays an equally important role in determining the overall efficiency of a particular display format. Hence, in order to effectively present the multiple sources of information contained in the cockpit instrument panel, "display formats should be constructed to take advantage of human information processing and pattern recognition capabilities" (Mahaffey, Horst, & Munson, 1986, p. 1514). In general, display design principles should be developed that capitalize on the basic principles of perception and of the perceptual-cognitive interface (Wickens, 1988).

Task Information Processing Characteristics

The successful application of any display design principle relies on an accurate assessment of the frequency and nature of the task information processing requirements, both of which can be determined by appropriate cognitive task analysis procedures (see also Rasmussen, 1986; Woods & Roth, 1986). Accordingly, multi-element display processing can be described by three subcategories of task requirements. These are requirements for the human operator: (1) to correctly integrate the displayed information that must be combined (information integration), (2) to engage in a sufficient degree of parallel processing so that critical displayed channels are not neglected even as information in other channels is attended (parallel independent processing), and (3) to allow, when needed, information to be correctly extracted from a single variable without being distorted or biased, and without attention being distracted by information in other channels (focused attention).

These subcategories can be characterized by their relative positions along a "task proximity continuum." Task proximity can be described in terms of the similarity in nature of the central processing requirements, or in terms of the similarity of mapping relations between the stimulus inputs and the subsequent response. Thus, task proximity is considered high when there is a many-to-one mapping of stimuli to a single response, and is considered low when there is a single, independent mapping of stimuli to respective responses. In cognitive terms, the degree of task proximity corresponds to the degree of relevance of two or more information processing channels to the unified goal of a task. That is, task proximity increases as the relevance between central processing channels increases.

Figure 1 depicts two end points of the task proximity continuum relevant to this study. Here, integration tasks are defined as high proximity tasks because multiple sources of information must be considered together (i.e., compared, computed, etc.) before the required response can be executed. In contrast, focused attention tasks are low proximity tasks because one needs to focus attention on a specific channel of information while "tuning out" other channels in order to execute the required response.

Consistent with this depiction of multi-element display processing, several studies have compared performance in tasks that require independent processing or focused attention with tasks that require the integration of multiple channels of information (e.g., Barnett & Wickens, 1988; Boles & Wickens, 1987; Carswell & Wickens, 1987; Goettl, Kramer, & Wickens, 1986; Peterson, Banks, & Gertman, 1981). A review of these studies reveals a pattern of crossover interactions between the task information processing requirements and the format of the display. It appears that the concept of proximity underlies both of these factors. The general finding is that integration task performance is best when the information to be integrated is displayed in close proximity (Polson, Wickens, Klapp, & Colle, 1989). Conversely, focused attention performance is best when the information to be extracted is somewhat more separated.

The Proximity Compatibility Principle

Collectively, the pattern of interactions between task and display proximity has both contributed to and supported the formulation of a theory-based

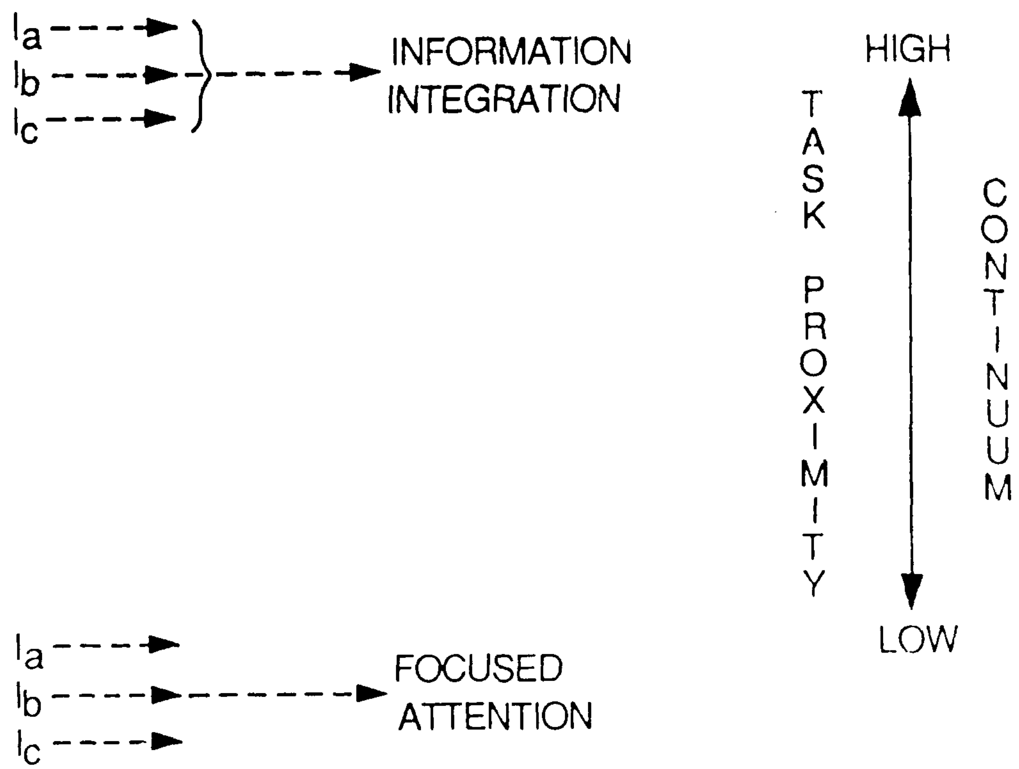

Figure 1. A task proximity continuum.

principle of display design referred to as the proximity compatibility principle (Barnett & Wickens, 1988; Boles & Wickens, 1987; Carswell & Wickens, 1987; Wickens, 1987). This principle attempts to relate the processing of the displayed information to the nature of the task information processing characteristics and asserts that tasks in which "close mental proximity" is required (i.e., information integration) will be best served by more proximate displays. On the other hand, tasks that require the independent processing of two or more variables, or the focusing of attention on one, while ignoring the others will be best served by more separate displays. Table 1 provides an outline of the relationship between display and task proximity as predicted by the compatibility of proximity principle.

Table 1

Compatible and Incompatible Matches of Task and Display Proximity

	High-Task Proximity	Low-Task Proximity
High-Display Proximity	COMPATIBLE	INCOMPATIBLE
Low-Display Proximity	INCOMPATIBLE	COMPATIBLE

Note. From Technical Report ARL-88-2/AHEL-88-1 by C. M. Carswell and C. D. Wickens, 1988, University of Illinois.

Display Proximity

Clearly, an objective specification of the defining characteristics of display proximity is essential for this principle to be applicable to "real world" interface design. The system designer must be able to interpret these characteristics in order to apply them within the compatibility of proximity framework (see Figure 1 and Table 1). Display proximity can be defined as the degree of nearness, or similarity, among display variables. At least three "macro dimensions," or ways of defining proximity can be specified. These dimensions are

(1) Physical metrics (e.g., closeness in space, similarity of color, similarity of acoustic parameters of speech). The guidelines to use functional grouping to position display elements, or to place graph legends close to the lines that they describe exemplify the relevance of this dimension.

(2) Similarity of representation (e.g., two digital displays are more similar to each other than a digit and a bar graph display; Boles & Wickens,

1987). The use of all digital or all analog displays to portray quantities that must be combined or compared exemplifies the relevance of this dimension.

(3) Object integrality (e.g., two dimensions combined into a single object are more proximate than two dimensions of two different objects; Carswell & Wickens, 1987; Barnett & Wickens, 1988). An example of this dimension is the aircraft attitude display indicator, which portrays both pitch and bank in a single object--the moving horizon--as these two dimensions must often be integrated to achieve level turns.

In this series of experiments, two display variables were brought to bear on the issue of designing a display to serve an aircraft stall monitoring task. These were space and color. According to the three-part classification of display proximity macro-dimensions described earlier, these two variables can be categorized as physical metrics. In the following sections we shall first describe the evidence that exists to suggest that each of these variables might underlie the proximity compatibility principle. Then we describe the general task scenario that provides the basis for each of the three experiments that follow.

Spatial Proximity

Space is an important and salient physical property that defines a channel of stimulus information (Wickens, 1984). Therefore, much of the research in the area of visual information processing has attempted to highlight the distribution of visual attention over space (e.g., Eriksen & Eriksen, 1974; Eriksen & Hoffman, 1972, 1973; Laberge, 1983; Podgorny & Shepard, 1983). A common finding is that "the ability to discriminate between elements in close spatial proximity is often severely limited" (Humphreys, 1981, p. 17). The result is either beneficial or detrimental depending upon the relations between stimuli and the required response.

Eriksen and Eriksen (1974), for example, had subjects respond to one of four centrally displayed letters with a corresponding left or right lever movement. The stimulus letter was presented by itself or was flanked by surrounding letters that were either compatible (indicating the same direction of motion) or incompatible (indicating the opposite direction of motion). Their results showed that close proximity of information within a spatial channel could produce benefits as well as costs to response time. Benefits were greatest when the flanking letters were identical to the central stimulus letter (i.e., a redundancy gain). Conversely, costs to response time were most pronounced when the flanking letters had implications for action that were incompatible with those of the central stimulus (i.e., response conflict). Most important to the present research is the finding that the magnitude of benefits and costs increased as the proximity of the flanking letters to the central letter increased. Thus, interference from incompatible letters was greatest in the focusing task when the letters were more physically proximal than when farther removed (Carswell & Wickens, 1988). Likewise, facilitation from compatible letters was greatest in this same task as the spatial proximity between the target and flanking letters increased. Similar results were found by Eriksen and Hoffman (1973).

Highlighting the general limits of focused attention, these results suggest that when displayed in sufficiently close proximity, two perceptual channels will be processed, even if only one is desired. In fact, Broadbent (1982) suggests that this "parallel" processing of information channels is more likely to take place when the information appears within 1^{o} of visual angle of a focused target. Eriksen and Eriksen's results point to the importance of spatial proximity in visual attention and suggest effects similar to those predicted by the proximity compatibility hypothesis. A general conclusion to be drawn is that "stimuli which need to be processed together should be close together in space, while stimuli that should be treated separately should be relatively distant" (Wickens & Flach, 1988, p. 119). The results from a majority of these experiments were interpreted within the framework of a "zoom lens" model of visual attention (e.g., Eriksen & Yeh, 1985; Eriksen & St. James, 1986). This model accounts for the perceptual characteristics that determine the focus of attentional resources, but does not consider the cognitive factors relevant to visual attention that are incorporated into the proximity compatibility principle.

While the importance of spatial proximity has been demonstrated with a variety of tasks including letter detection (e.g., Eriksen & Eriksen, 1974) and feature analysis (e.g., Humphreys, 1981), these experiments can be generalized to real-world settings only to a limited degree. One reason for this limitation is that the nature of the stimuli (e.g., letter targets) and the presentation format (i.e., tachistoscopic views) used in the majority of these experiments has little relation to the cockpit setting where visual scanning and sampling of multiple, complex information channels takes place.

An exception is a study by Holahan, Culler, and Wilcox (1978) in which the number, color, and spatial proximity of distractor signs to a target stop sign were varied in a simulated traffic environment with free field viewing (i.e., not tachistoscopically controlled). The subjects' task was to scan the display for a target stop sign and to identify its presence or absence as quickly as possible. The results indicated that the subjects' ability to discriminate the target from the distractors was primarily dictated by spatial proximity. When distractors were displayed in close spatial proximity to the target, response times in all conditions were high; but when displayed at a distance, the number and color of the distractors differentially affected response time.

Although Holahan et al. (1978) provide a practical application of physical display manipulations, these results are still not entirely generalizable to the cockpit setting. The reason for this assertion is that the set of tasks carried out by aircraft pilots varies over the task proximity continuum. However, measures of the spatial extent of attention have most often been derived from low proximity (focused attention or independent search) task performance, while performance in high proximity (information integration) tasks has rarely been considered or contrasted.

Only recently has an effort been made to examine the relationship between the display proximity of multiple stimulus inputs and the corresponding mental proximity of the task central processing characteristics. In particular, three experiments have examined the role of spatial proximity in the context of the compatibility of proximity paradigm.

In a simulated airborne decision experiment carried out by Barnett and Wickens (1988), spatial proximity was manipulated by presenting cues either at separate locations on the screen or in the center of the screen. Subjects were asked to compute the information of four cues and then to integrate their information values in order to determine whether or not to abort a current mission. The results of this study did not indicate a significant benefit in integration performance for the condition of close proximity in space (centered on the screen versus four distinct locations).

A study by Pamperin and Wickens (1987) also manipulated the spatial proximity of information relevant to a task. In this study, subjects were required to monitor a display of dials while responding to a two-letter probe that appeared either in the center of the array of dials or in the periphery of the display. Thus, the two levels of visual presentation provided a contrast between high (center) and low spatial proximity (peripheral). An integration task required subjects to determine if two vectors (as designated by a two-letter probe) in the alphabetically labeled dials were rotating in the same direction. A dual-task condition presented the same two elements of display information (rotating vectors and a two-letter probe), but required independent decisions and responses to be made on each display element.

Their results indicated that the spatially proximate (central) condition provided slightly faster response times than the wider spatial dispersion (peripheral) for the integration task as well as the dual-task condition. In this case, spatial proximity did not differentially affect the high and low proximity tasks. However, it is possible that the peripheral condition impaired performance in both tasks because of the difficulty of processing physically disparate stimuli.

Finally, a study by Harwood, Wickens, Kramer, Clay, and Liu (1986) also examined the effects of spatial proximity on information integration. In their study, subjects monitored two imaginary teams that were competing in their search for a lost treasure and made decisions on how to allocate resources within and between teams. Information about the teams varied along four different dimensions or attributes. Spatial proximity was varied between two levels: attributes of each team were grouped on the display in a separate spatial location (low spatial proximity) or all attributes of both teams were grouped together in one spatial location (high spatial proximity).

Their results suggested the importance of spatial organization. However, the form of grouping the eight elements composing the two teams improved the ability of subjects to make judgments based on the integration of information both within a team and across teams. Thus, spatial separation created an organization that facilitated information integration processing, but not in a way that directly supported the principle.

Uniformly, results from these three studies do not suggest support for the proximity compatibility principle when proximity is defined by physical space. However, further examination is warranted in a paradigm that includes an important element, present in most "real-world" display contexts, but missing in the experimental tests just reviewed. That element is irrelevant clutter. This was an important component of these experiments. In addition, with the exception of the Pamperin and Wickens study, none of these studies fully manipulated both spatial and task proximity levels in a way that would allow a strong test of the hypothesis.

Color Proximity

The guideline of close spatial proximity for integration tasks is intuitively plausible, and while not supported by research in clutter-free environments, it is anticipated to be relevant when clutter is present. However, cockpit design limitations make it nearly impossible to adhere to this guideline. One reason for this constraint is that different combinations of displayed information must be integrated at varying points in time. For example, airspeed must be integrated with a certain set of variables to determine a navigational status, while at some other time, it must be integrated with a different set of variables to determine aircraft stability. The problem then arises concerning the optimal location of the airspeed indicator.

A possible solution is to reinforce, supplement, or augment spatial proximity factors with variations in color coding, because these will also influence the perception of proximity (Kramer et al., 1986). Principles and uses of color in applied environments have been described in other studies (e.g., Boff & Lincoln, 1988; Silverstein, 1987). Many of these studies refer to issues such as the benefits of redundant color in target search or the advisable number of colors for coding. Relevant to the present experimental paradigm, it is important to examine whether color behaves as a physical variable that conforms to the compatibility of proximity principle. If so, integration performance should be enhanced when relevant display elements are presented in the same or similar color, while focused attention performance should be enhanced when the relevant display elements are presented in distinctly different colors. Thus, a second goal of this study was to examine the generality of the compatibility of proximity hypothesis regarding this second physical dimension of stimulus information--color.

Recent technological advances have made multicolor displays a realistic design option. As a result, an extensive effort has been undertaken to determine the relative advantages and disadvantages of color displays over monochromatic displays. Although little objective evidence exists to support color advantages (Hale & Billmayer, 1988), this has not interrupted the introduction of such displays in many commercial aircraft such as the Boeing 757, 767, and the European Airbus A310, as well as military aircraft including the F-14, F-18, F-15, A-7, ALPHA-JET, Mirage 2000, and Tornado. Presumably, the implementation of these displays has been driven somewhat by studies showing widespread pilot preference for color displays (see Aretz & Calhoun, 1982; Reising & Calhoun, 1982;) and by advancements in visual display technology (see Brindle & Mulley, 1984). However, it is uniformly agreed upon that designing with "aesthetic overindulgence" should be avoided (Martin, 1984), and that sound human factors principles should be adhered to in designing color displays (Godfrey, 1982).

Most of the research conducted on CRT displays has centered around the objective physical and/or subjective psychophysical legibility and image quality measurements in the assessment of both monochrome (e.g., Buffett, 1986) and multicolor CRT displays (e.g., Brindle & Mulley, 1984). Most often, these assessments are dictated by the available display technology and the environmental conditions under which they are taken. However, the perceptual/cognitive factors of display design (see Wickens, 1987, for a review) have rarely been considered, and only recently have researchers argued that color formats should be task dependent, that is, based on the appropriate set of tasks to which they are proposed to support performance (e.g., Berggrund, Derefeldt, Hedin, Marmolin, 1984; Hudson, 1984; Narborough-Hall, 1985; Santucci, Menu, & Amalberti, 1984; Snyder, 1984; Wickens, 1987).

Previous studies addressing the more cognitive aspects of color processing have provided evidence to suggest that the color-task relationship falls within the domain of the proximity compatibility principle. In general, color has been shown to strongly affect the relationships between attributes (i.e., the perceptual organization) of a display. Gestalt principles of perception (see Wertheimer, 1958) suggest that visual search is affected by perceptual organization (Taylor, 1985). Hence, visual search for items in a target color are organized according to Gestalt principles of proximity and similarity (Bundesen & Pedersen, 1983). Same-colored stimuli, for example, will more likely lead to a perception of proximity than will separately colored stimuli. In addition, same-colored stimuli will more likely be remembered together than those of different colors (Moar, 1977). Accordingly, color can be used to group informational sources that need to be integrated, even when these sources are displayed in physically separate locations (Kramer et al., 1986). Clearly, to the extent that color display variables can restrict the focus of attention to relevant information channels, performance will benefit. Thus color can also be used to "unite" or "chunk" relevant display elements and partition them away from other irrelevant stimuli in the visual field (Reising, Emerson, & Aretz, 1984).

While it is clear that color can be useful in grouping relevant information within and across displays, relatively few studies have investigated color processing efficiency as a function of the task requirements. Those that have are in agreement that research results concerning performance benefits gained from color are somewhat variable and task dependent (Stokes & Wickens, 1988). After reviewing several studies investigating the usefulness of color-coded information, Christ (1975) concluded that results should be analyzed and compared based on the task that subjects had to perform when using color-coded information. In particular, he distinguishes between search tasks and identification tasks.

An analysis of empirical data regarding color in search task performance has shown color-task relationships consistent with the proximity compatibility hypothesis. For instance, color coding is generally found to improve search task performance when the target is uniquely coded from non-targets. Likewise, performance is best when the target-background color difference is large (Carter, 1982). However, this facilitation is less likely to occur when the target is of the same color as other stimuli in the visual field. This fact is highlighted in a study by Holahan, Culler, and Wilcox (1978) described in the previous section, in which the effects of both color and spatial proximity were studied using a visual search task in a simulated traffic environment. When distractors were displayed in close spatial proximity to the target, response times in all conditions were high; but when displayed at a distance, the dominant cause of increased response time to identify a red target stop sign was the presence of some distractors of the same color. Thus, coding all display elements in the same or similar color reduced the subject's ability to discriminate (i.e., focus attention on) the target effectively.

Although many studies have examined color effects in search task performance (Christ, 1975), it is important to note that real-world tasks generally do not deal with only simple categories, such as search and identification (Narborough-Hall, 1985). Instead, tasks are often combined and cannot easily be categorized as such. Hence, it is important to reveal whether the relations between color on a display and the higher-order mental representations of a task formed by the observer support the proximity compatibility hypothesis. Is it, as Narborough-Hall (1985) suggests, more difficult to make distinctions (focus attention) between stimuli when they are coded by the same color? Conversely, do stimuli coded by different colors inhibit complex comparisons or computations (information integration) across them? Here research results are more scarce.

Most relevant to the current research is a study by Harwood et al. (1986) described in the previous section, in which color coding and spatial position levels were manipulated to emphasize the identification of relationships between attributes of teams. Space and color were used two ways: (1) to emphasize attribute relationships within teams, and (2) to emphasize attribute relationships between teams. Results indicated that integration performance was best when the teams were displayed in proximity; that is, either spatially proximal or similar in color. This relationship was upheld whether comparisons were made within or between attributes.

A summary of the results of this experiment reveals that both color and spatial proximity improved performance in this integration task although the beneficial effects of these two dimensions were not additive. Thus either code alone, spatial proximity or common color, produced the same reduction in reaction time as did both in combination. This suggests that poor "grouping" on one dimension (space or color) could be compensated for by emphasizing the grouping of relevant information on the other dimension. However, the type of color code, colored by attributes or by teams, did not discriminate between task comparisons of attributes or teams and thus did not directly support the principle. In addition, the investigators did not examine how these variables affected the ability to focus attention on the displayed information.

Clearly, there is little empirical evidence to provide insight into the role of color in complex task performance. What evidence there is, however, appears to suggest that display elements coded in the same or similar color facilitate the integration or grouping of those elements relevant to a common task (Harwood et al., 1986), and effectively degrade performance when one must focus attention on a specific source of information (Holahan et al., 1978). Although these results partially support the compatibility of proximity hypothesis, more objective evidence is needed to confirm the relationship between the perception of color proximity and the proximity of the cognitive task demands.

EXPERIMENTAL OBJECTIVES

Unfortunately, few of the studies investigating the usefulness of color coding have simulated the real-world conditions confronting the pilot of an advanced aircraft (Stokes & Wickens, 1988). In addition, many of the studies investigating the effects of display proximity on information integration have used displays consisting of simple graphics or geometric objects in very hypothetical task contexts.

With these limitations in mind, the purpose of these three experiments was (1) to assess the extent to which manipulations of spatial closeness and color similarity among displayed sources of information adhered to the proximity compatibility principle, and (2) to test this principle in a more realistic task context by displaying three parameters pertaining to the likelihood of aircraft stall, using the simulated dynamics of a light aircraft. In the first two experiments, irrelevant visual clutter is present on the display. In Experiment 3, clutter is absent. Experiments 1 and 2 contrast monochrome displays with displays in which the three relevant indicators are multicolored and commonly colored, respectively. Experiment 3 directly compares these two color coding schemes. All three experiments impose the same general task requirements that call for both information integration and periodic focused attention check reading. In this way, the effectiveness of the proximity compatibility principle with regard to space and color can be assessed.

The present experiments used a display to present three critical variables that interact and determine the probability that a fixed-wing aircraft will stall: airspeed, bank, and flap setting. The stall task was chosen not only because it is prototypical of an information integration task in which the value of several variables must be considered together to assess the danger of stalling, but also because stability is a critical concern in the cockpit of either fixed-wing or rotary-wing aircraft. Thus, the task at once meets the criterion of a valid test of the compatibility of proximity principle and obtains a degree of real-world validity.

Based on the compatibility of proximity principle, an interaction between display format and task type was expected. Display formats that emphasized the perceptual grouping of relevant information (through common color or spatial proximity) were expected to best support integration performance. Conversely, those formats that emphasized the distinct identities of the information (through separate colors or spatial separation) were expected to best support focused attention performance.

EXPERIMENT 1: UNIQUE COLORS

Method

Procedure

Displays were viewed on an IBM® color monitor driven by an IBM-XT with an enhanced graphics adapter (EGA) board. Subjects were seated directly in front of the screen at a distance of 50 centimeters and viewed a set of five dial-like indicators displayed for 1.5 seconds. Three of these indicators contained relevant information pertaining to airspeed, bank, and flap settings. Each indicator reading was standardized within a range of _0_ (minimum) to _10_ (maximum). Display clutter was incorporated in all conditions by presenting two indicators (which showed no values) along with the three relevant indicators. Immediately after display termination, subjects were prompted with one of two task probes using the retrospective probe technique developed by Carswell and Wickens (1987) and used by Barnett and Wickens (1988) in their study of object-display integration.

Figure 2 shows a schematic representation of the two task operations required in the experiment. The subjects' primary task was to integrate the values of the three displayed indicators to determine the probability of a stall on a scale from 1 to 10. The correct formula for stall likelihood was approximately based on an analytic representation of the dynamics of a light aircraft, bearing a positive relation to bank and a negative relation to flaps and airspeed. Airspeed was the dominant factor in the equation, and interactive components were included such that nonoptimal settings of flaps and bank had greater impact at low airspeeds. On the average, subjects were prompted with the integration task probe 75 percent of the time. However, the other 25 percent of the task probes required the subjects to recall the specific value of one of the indicators (focused attention) on a scale from 0 to 10. Task type was signaled to the subjects by the words "stall," "airspeed," "bank," or "flaps," displayed on the screen immediately after termination of the information display. Responses to all probes were entered using the keyboard, and instructions were given to respond as quickly and as accurately as possible. Subjects were given feedback showing their response as well as the correct response for each trial. Three seconds after their response, subjects were automatically cued for a new trial with the presentation of a centrally located fixation cross.

Information could be displayed in any of six formats. Spatial separation among the relevant indicators was varied at three levels: close, intermediate, and distant (see a, b, and c on Figure 3.) Within each of the three levels, color proximity was varied between two levels: a color-coding scheme in which the relevant indicators were uniquely colored or contrasted with a monochrome display of all five indicators. All indicators in both the monochrome and color conditions were presented against a black background. In the color display, the three relevant indicators were presented in separate colors to emphasize the distinction among the relevant variables. The airspeed indicator was white, while the bank and flap indicators were magenta and blue, respectively. In addition, both clutter indicators were white. Thus, while there is a distinction in color among the three relevant indicators, there exists no distinction in color between the airspeed indicator and the irrelevant clutter. (This coding scheme could not be avoided because only three colors were available on the display in addition to the black background.)

Design

The experiment was conducted in a 2 (tasks) x 2 (color formats) x 3 (spatial formats) within-subjects factorial design. Subjects participated in one 1-hour session during which they were exposed to all combinations of factors (i.e., all six display formats) for a total of 180 trials. Spatial separation levels were counterbalanced over 6 blocks of 30 trials each in order to negate any learning effects. The first three blocks consisted of distant, close, and intermediate spatial separation levels, respectively. This order was reversed for the last three blocks. Within each block, color coding levels (i.e., separate or monochrome) changed every five trials, while the task probes (integration or focused attention) were randomly presented. Latency and accuracy measures were evaluated as indices of performance.

All subjects were given explicit instructions concerning the method for determining the probability of stall, followed by a series of 10 practice trials

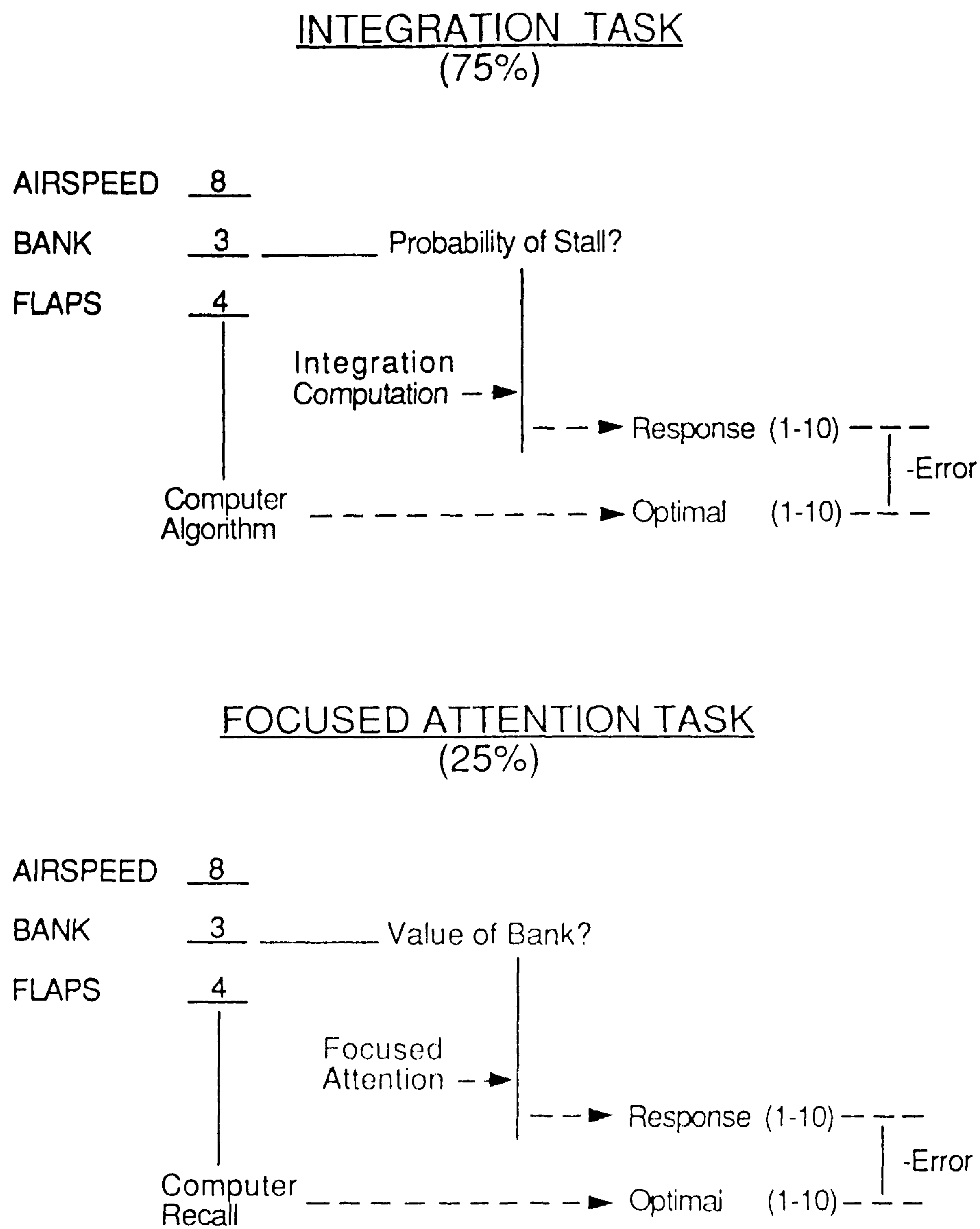

Figure 2. Logic of displays, probe questions, and responses for the integration task (top) and the focused attention task (bottom).

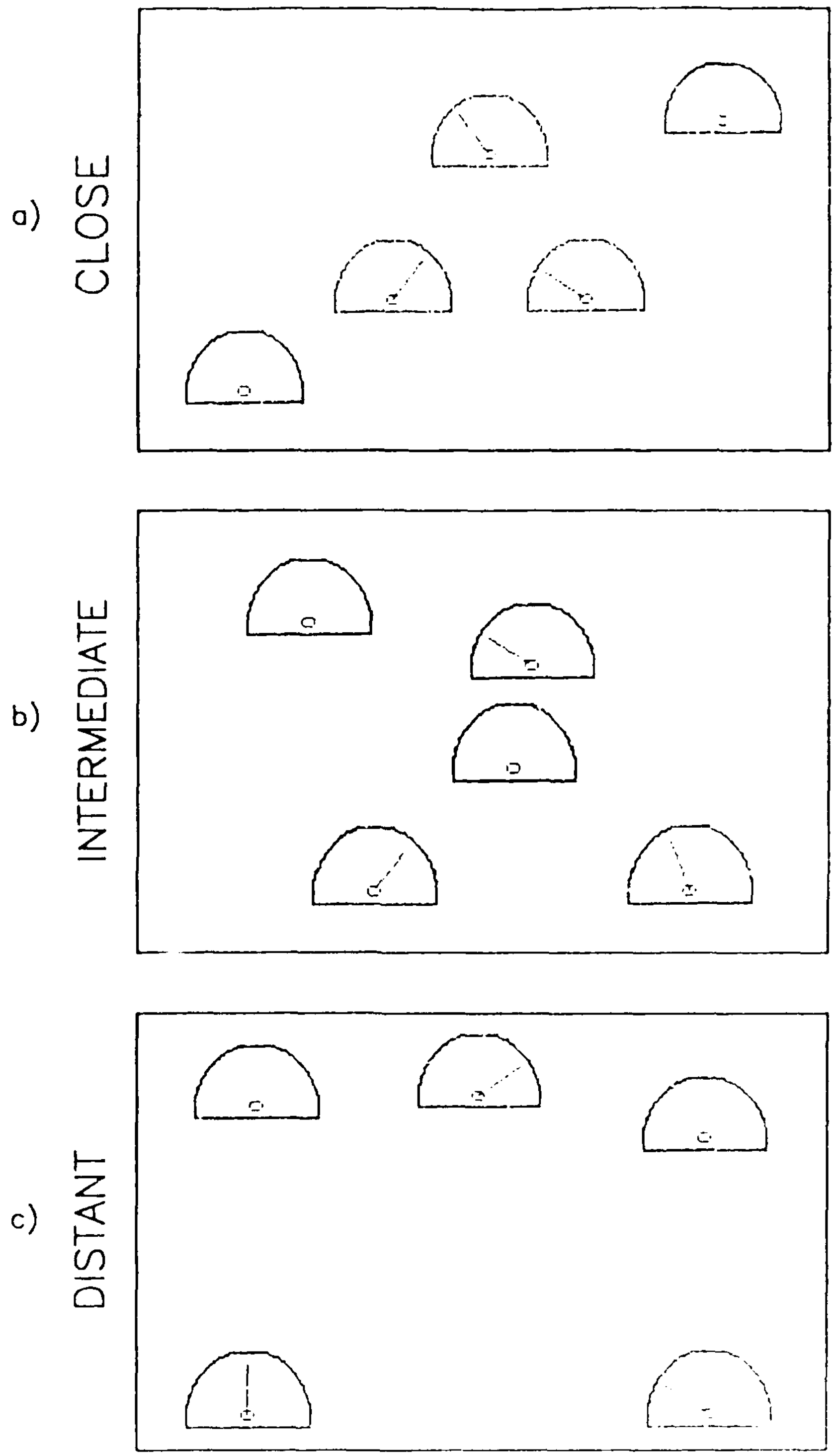

Figure 3. Three display spatial configurations used in Experiments 1 through 3. (In Experiment 3, the two irrelevant clutter indicators were removed.)

to ensure a proper conceptual knowledge of the program formula. Only after a demonstration of this knowledge were subjects allowed to continue with the actual experimental trials. The instructions provided (1) a description of the three variables, (2) a conceptual understanding of the way the three variables interact to determine the likelihood of a stall, and (3) the relative weighting or importance of the three variables in the formula.

Subjects

Nine subjects (undergraduates attending the University of Illinois) were paid to participate in the experiment. All subjects had normal or corrected-to-normal vision and were screened for their ability to perceive the actual colors used in the experiment (by self report). The subjects had no flight experience.

Data Analysis

Analysis of the results of Experiments 1, 2, and 3 will be organized in terms of the effects on performance of the two primary experimental variables: color and space. Within each, we will discriminate the effect of that variable on the different tasks (integration and focused attention), as reflected in both accuracy and latency. Accuracy (expressed in error magnitude) was measured by the average absolute error between the subject's response and the correct response for each trial.

Results

At the outset, it should be noted that subjects were generally able to use the stall formula effectively. Across all probes, the mean correlation between optimal and actual integration values was 0.72. Airspeed values contributed the greatest weight in accounting for variance of the response, while bank and flaps contributed smaller, but still significant weights, respectively.

Color Proximity

Analysis of the error data failed to reveal any significant differences between the two color formats for performance on the integration task. In addition, no main effect was found between the color schemes for focused attention performance. However, a significant Color x Indicator interaction showed that the separate color code improved the accuracy (reduced error) of focused attention recall of the bank values and degraded recall of the airspeed indicator (which was the same color code as the noise elements) relative to the monochrome display, $\underline{F}(2,16) = 6.33$, $\underline{p}<.01$. Recall accuracy of the flap values did not differ between the two color formats.

Analysis of the reaction time data revealed important differences between the two color schemes. Figure 4 plots the reaction time data for the two types of tasks (integration and focused attention) as a function of the color coding schemes. The data are averaged across levels of spatial separation, while the focused attention data are averaged across indicator type. The Display x

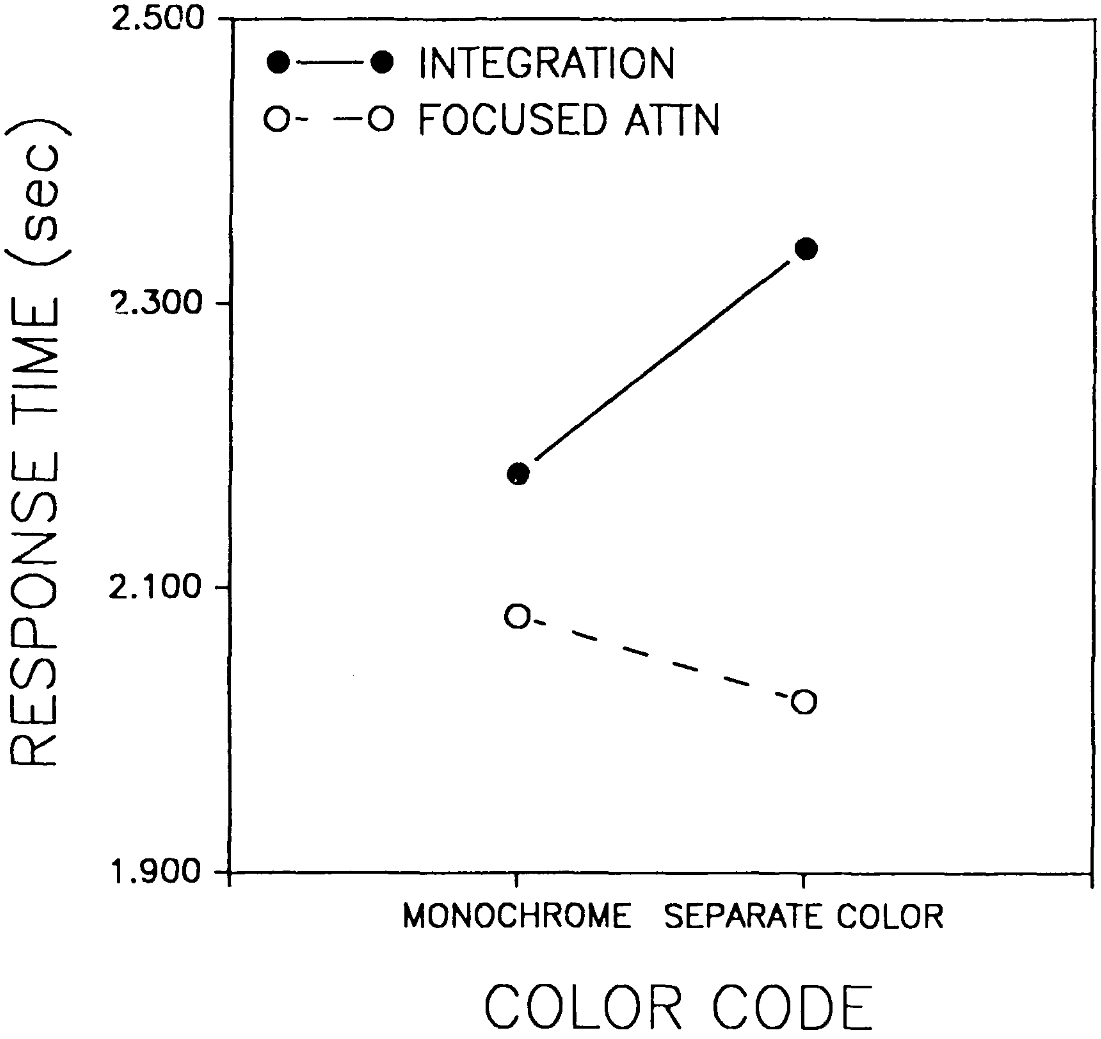

Figure 4. Effect of color code and task type on response time (Experiment 1). (The focused attention data are averaged over indicators, and all data are averaged over spatial separation.)

Task interaction for reaction time data did not reach a level of significance. However, the trend is in a direction that supports the proximity compatibility hypothesis. Although the use of a separate color code significantly degraded integration performance ($F[1,8] = 12.32$, $p<.01$), focused attention performance showed only a slight (but nonsignificant) improvement relative to the monochrome condition. An analysis of the effects of color code and indicator type provides some evidence of why the separate color code did not produce a significant improvement in response time for focused attention recall. This trend, shown in Figure 5, is of the same general form as that observed with the error data. It reveals that the distinct colors facilitated recall of the bank and flaps indicators, but at the expense of the processing and recall of the airspeed indicator, which was the same color as the irrelevant clutter. Thus, the overall improvement in focused attention response time (and reduced error) resulting from the use of the separate colors was offset by the delayed response time (and increased error) in airspeed recall.

Spatial Proximity

The manipulation of spatial proximity showed that close spatial proximity did not foster integration. In fact, both the close and intermediate spatial configurations disrupted accuracy of the integration task relative to the distant proximity condition in which integration performance was significantly more accurate, $F(2,16) = 6.37$, $p<.05$. Spatial proximity also had an effect on focused attention accuracy, $F(2,16) = 10.63$, $p<.01$; but this effect was not monotonic. The Tukey test for mean differences ($p<.05$) revealed that both the distant and the close proximity conditions provided significantly higher accuracy than did the intermediate proximity condition. Thus, considering the results of both tasks together, performance was best when the relevant information was widely separated and worst when they were presented at intermediate levels of dispersion.

Reaction time data failed to reveal significant effects for manipulations of spatial proximity for either focused attention or integration performance. The only significant Color x Space interaction occurred in the reaction time data for integration performance, $F(2,16) = 7.42$, $p<.01$. This interaction indicated that the separate color code hurt integration performance the most at the widest spatial separation.

An analysis of the accuracy measures for recall of the three indicators revealed a significant main effect, $F(2,16) = 6.81$, $p<.01$. The Tukey test ($p<.05$) showed that recall of the airspeed gauge was significantly more accurate than recall of the other two gauges, showing a nonsignificant trend in favor of bank over flap setting recall.

Reaction time data provided a main effect for indicator recall, $F(2,16) = 7.90$, $p<.01$, which was identical to the results for the accuracy measure. The Tukey test ($p<.05$) revealed that recall of the airspeed value was significantly faster than that of the other two gauges. As with accuracy, the latter two showed a nonsignificant trend in favor of bank over flap setting recall speed.

The results from manipulations of color proximity were consistent with the compatibility of proximity principle. When performance was measured by response time, the use of a separate color code to impart a "psychological distance" between information channels was observed to produce benefits for the focused attention recall of those variables that were uniquely colored in the display, and costs to information integration performance relative to the monochrome condition. The error data revealed that these effects were not the results of a speed-accuracy tradeoff and further pointed to the disruptive effect of color proximity on focused attention when this proximity grouped relevant (airspeed) and irrelevant (clutter) information into the same perceptual category.

In contrast, results from manipulations of spatial proximity levels, when significant, did not appear to follow the proximity compatibility principle. Overall, accuracy of performance was best when spatial separation was at its greatest and poorest when spatial separation was presented at intermediate levels. Furthermore, this pattern was not modulated by task type as predicted by the hypothesis. This effect seems to be because of the confounding of distance with clutter position and will be explored further in Experiment 2.

It is important to note that the formula used was successfully learned by the subjects. The magnitude of errors as well as reaction times shown in the focused attention recall trials (across the three indicators) was inversely related to the weight assigned to each variable in the formula. This weighting, however, was distorted when the least important indicators were highlighted by the salient color coding, as shown in Figure 5.

EXPERIMENT 2: COMMON COLOR

As in Experiment 1, the second experiment tested the compatibility of proximity principle when proximity was manipulated through the use of spatial positioning and color coding. The purpose of Experiment 2 was (1) to discern the effects of a common color code to the three relevant indicators on the same set of tasks (integration and focused attention), and (2) to replicate the effects obtained in the previous experiment concerning spatial proximity. With regard to manipulations of color proximity, this design was predicted to produce results opposite to those obtained in Experiment 1. That is, the use of a common color code to impart a "psychological proximity" between all relevant information channels was expected to produce benefits for information integration performance and costs to focused attention performance relative to the monochrome condition.

Method

Design and Procedure

The design and procedure for Experiment 2 were identical to those used in Experiment 1 with the exception of the color coding scheme. In this experiment, a common color scheme which coded the three relevant variables in the same color (magenta) and separated them from the clutter (white), was contrasted

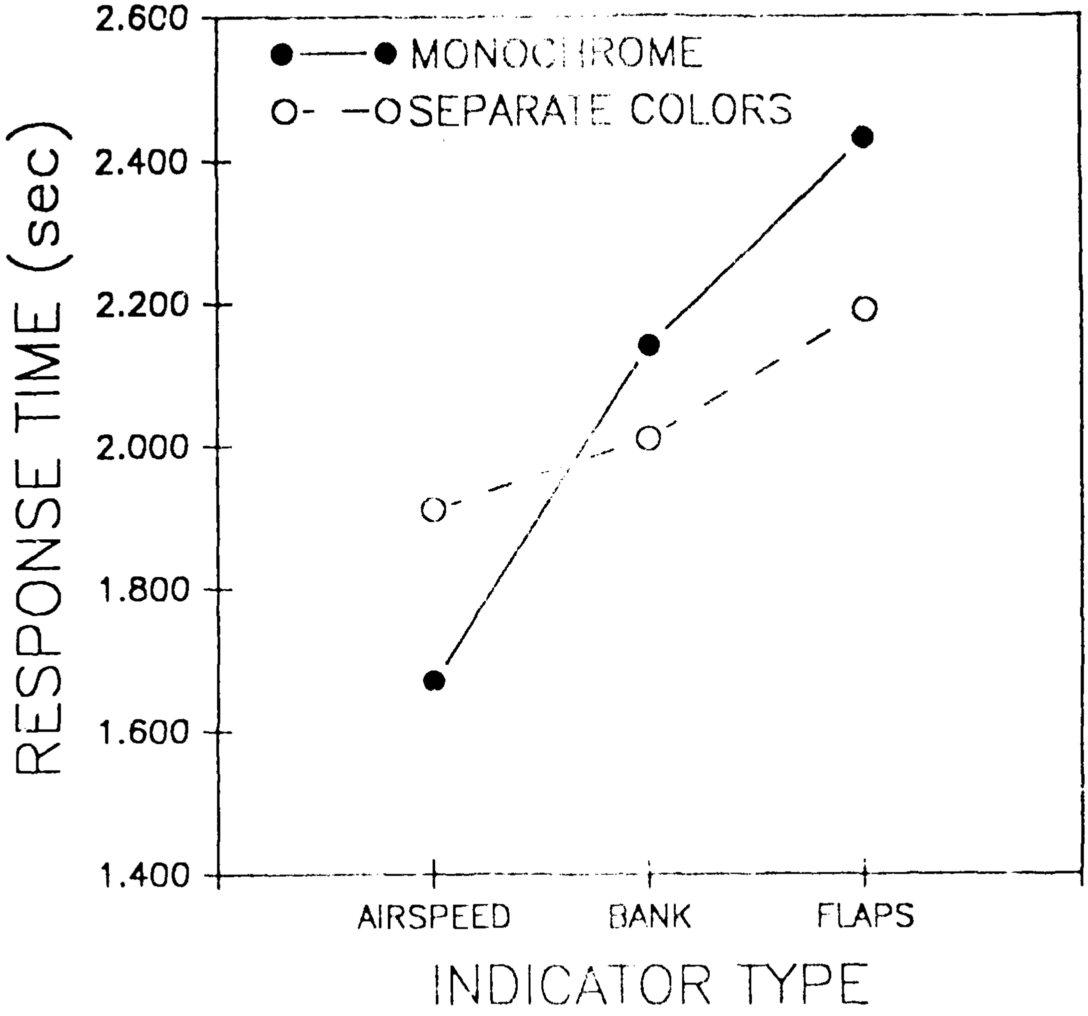

Figure 5. Effect [illegible] on response time (Exp[illegible])

with a monochrome condition, identical to that used in the previous experiment. As in Experiment 1, spatial proximity was varied among three levels: close, intermediate, or distant (see a, b, and c on Figure 3.)

Subjects

Eight subjects (undergraduates attending the University of Illinois) were paid to participate in the experiment. All subjects had normal or corrected-to-normal vision and were screened for their ability to perceive the actual color used in the experiment (by self report). The subjects had no flight experience.

Results

Color Proximity

Figure 6 plots the error data for the two types of tasks (integration and focused attention) as a function of the color coding schemes. A significant Display x Task interaction for accuracy data supports the proximity compatibility hypothesis, $F(1,7) = 9.27$, $p<.05$. That is, the use of a common color to code the relevant indicators improved integration task performance, while focused attention accuracy was significantly degraded relative to the monochrome condition. Separate comparisons revealed that the degradation of focused attention performance was reliable ($F[1,7] = 5.77$, $p<.05$), but the improvement in integration performance was not ($F[1,7] = 3.21$, not significant [ns]).

Analysis of the reaction time data failed to reveal significant effects between the two color formats for either focused attention or integration performance.

Spatial Proximity

Analysis of the accuracy data for the integration task revealed a main effect of spatial proximity, $F(2,14) = 5.25$, $p<.05$. Replicating the trend observed in Experiment 1, the Tukey test for mean differences revealed that the distant proximity condition provided superior integration accuracy over the intermediate condition, while performance in the close proximity condition was of intermediate accuracy and was not significantly different from either of these two levels ($p<.05$). The accuracy data on focused attention trials failed to show a significant effect across levels of spatial proximity.

Reaction time data failed to show a significant effect on integration performance. However, focused attention trials showed a nonsignificant trend in which the distant and close proximity conditions provided somewhat faster response times than the intermediate condition ($F[2,14] = 2.92$, ns). As with Experiment 1, the manipulation of spatial proximity did not show the Display x Task interaction predicted by the compatibility of proximity principle.

The only significant Color x Space interaction was found in the accuracy data for focused attention performance ($F[2,14] = 5.90$, $p<.05$). This pattern of results indicated that the common color code degraded focused

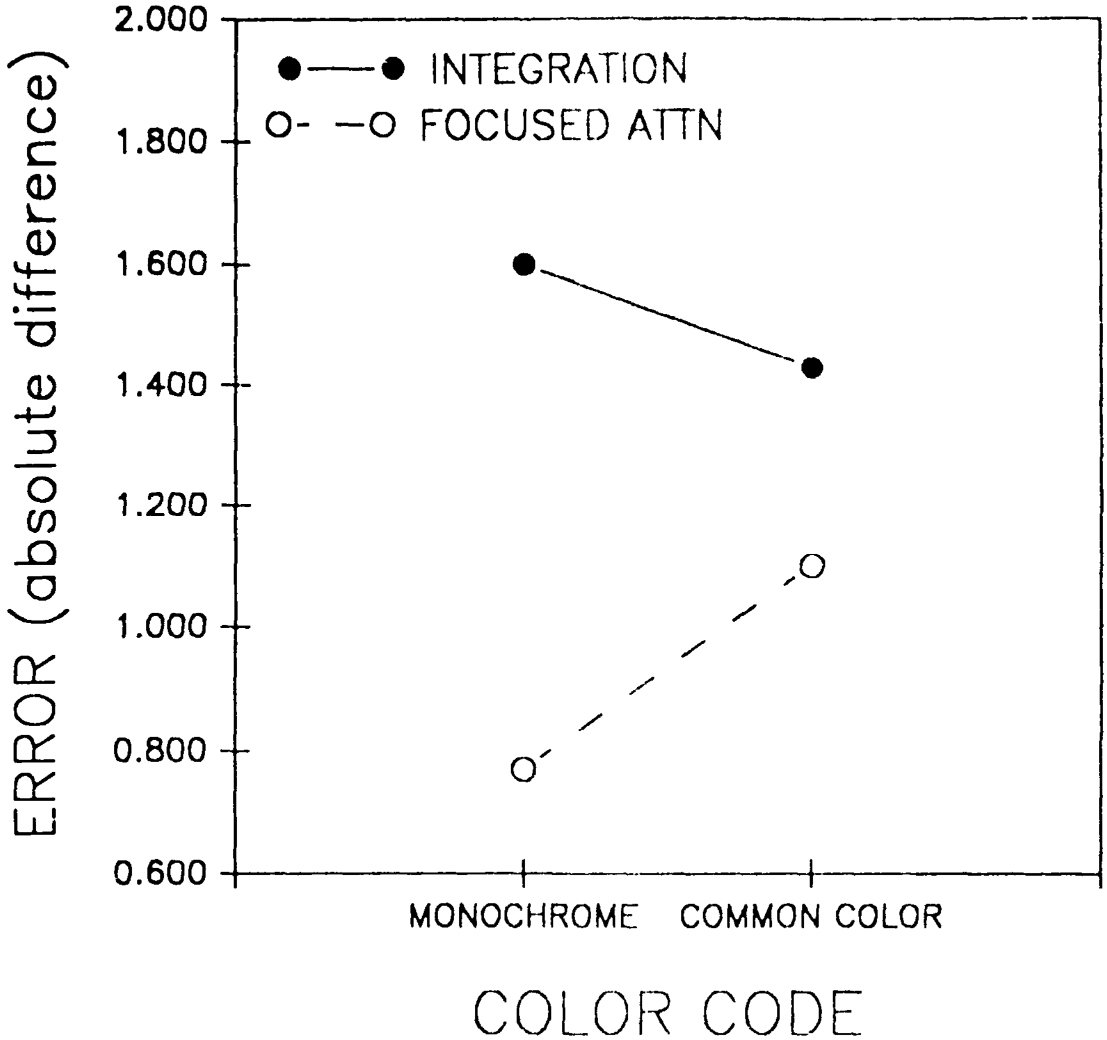

Figure 6. Effect of color code and task type on error (Experiment 2).

attention accuracy most at the close and intermediate spatial proximity levels while it did not affect performance at the widest spatial separation.

As with Experiment 1, subjects provided evidence of correctly weighting the three variables in accordance with the computer's stall formula. This weighting was revealed by a main effect of indicator type for accuracy, $F(2,14) = 13.0$, $p<.001$. The Tukey analysis shows that the most important indicator (airspeed) was responded to with the greatest accuracy followed by the bank and flap indicators, respectively ($p<.05$).

Reaction time data also revealed a main effect for indicator type ($F[2,14] = 7.56$, $p<.01$) showing that the airspeed indicator was responded to the fastest followed by the bank and flap indicators, respectively (Tukey, $p<.05$).

Discussion

Results from the manipulations of color proximity levels were again consistent with the compatibility of proximity principle. That is, the use of a common color code to impart a "psychological proximity" between information channels was observed to produce a nonsignificant benefit for information integration accuracy and significant costs to focused attention accuracy relative to the monochrome condition.

Replicating the results of Experiment 1, the significant effects of spatial proximity manipulations did not follow the proximity compatibility principle. Overall, accuracy was best when spatial separation was at its greatest and poorest when spatial separation was presented at intermediate levels. Furthermore, spatial proximity failed to interact with task type in a way that would be consistent with the principle.

The non-monotonic effect of spatial proximity revealed in both Experiment 1 and 2 is shown in Figure 7, which presents error averaged across focused attention trials and integration trials. The figure shows generally best performance (for both integration and focused attention) when the relevant indicators were separated the most and poorest performance (for both integration and focused attention) at intermediate levels of separation. While contrary to the view that pure spatial proximity between relevant indicators is a critical variable, these results are better understood by considering the position of the three stall indicators relative to the clutter indicators in each of the spatial proximity conditions. These distances were measured on the display screens depicted on a, b, and c of Figure 3.

For the close spatial proximity condition (see Figure 3a), the average distance between the relevant and irrelevant (clutter) indicators is 9.54 cm. For the intermediate spatial condition (see Figure 3b), the average distance between the relevant and irrelevant (clutter) indicators is 7.67 cm. However, for the distant proximity condition (see Figure 3c), the average distance between the relevant and irrelevant (clutter) indicators is 11.29 cm. Figure 8 plots these distances in the solid lines. Thus, for intermediate levels of spatial proximity, where clutter is closest to the relevant indicators, performance suffers to the greatest extent over both tasks. However, for distant levels, where clutter is least proximal, performance is superior for both tasks. The striking resemblance between the pattern of performance in Figure 7 and the

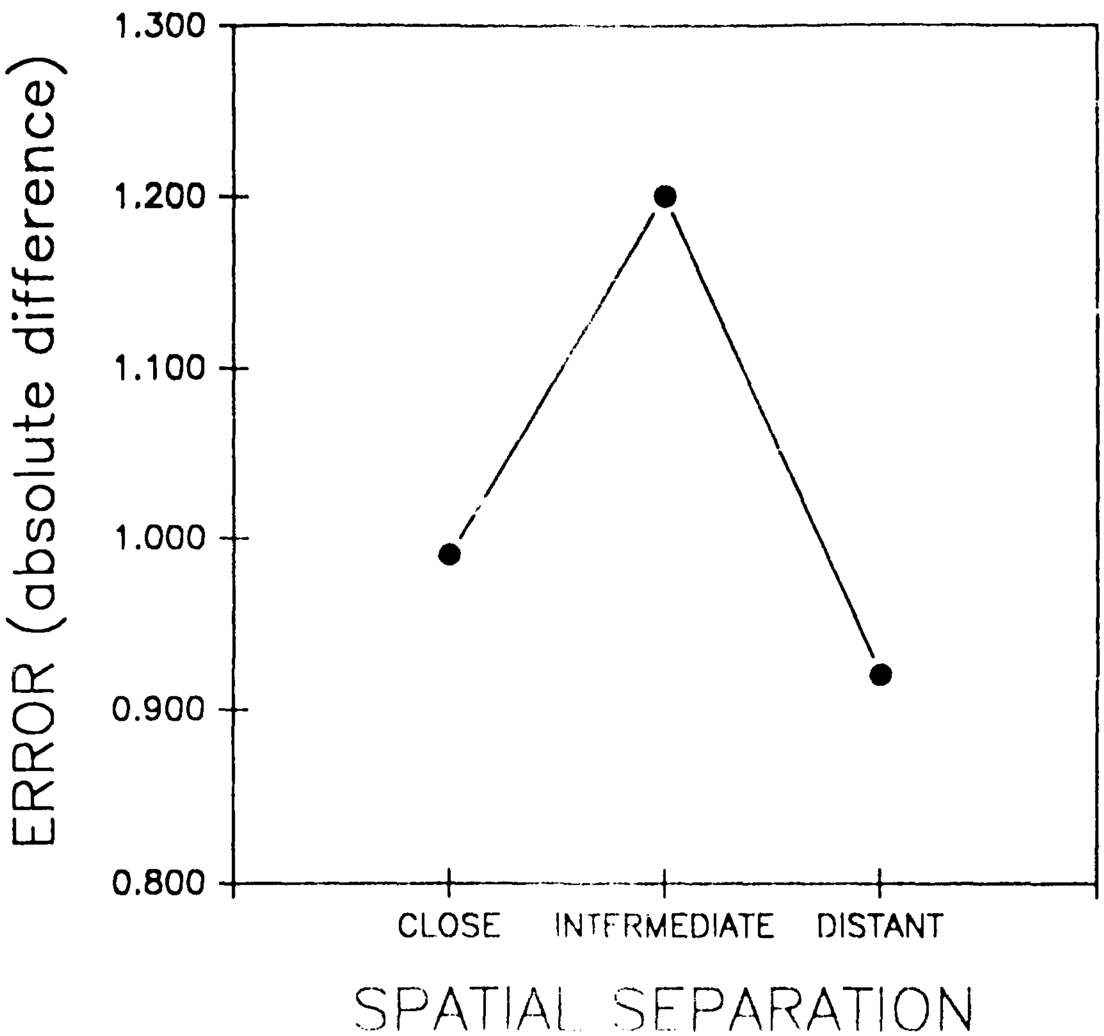

Figure 7. Effect of spatial separation on error (Experiment 2). (The data are averaged over task type and color code.)

pattern of distances in Figure 8 is highlighted by the dashed lines in Figure 8, which plot the accuracy of performance (1-error). The similarity between these two lines in Figure 8 suggests a distinct causal relationship between proximity to irrelevant clutter and performance on both focused attention and integration tasks.

EXPERIMENT 3: NO CLUTTER

As in the two previous experiments, this experiment tested the compatibility of proximity principle when proximity was manipulated through the use of spatial positioning and color coding. However, the purpose of this experiment was (1) to directly compare the common color display to the separate color display for the same set of tasks (integration and focused attention), and (2) to investigate the effects of spatial proximity in the absence of irrelevant display clutter. With regard to manipulations of color proximity, the use of a common color code to impart a "psychological proximity" between information channels was expected to produce benefits for information integration performance and costs to focused attention performance relative to the separate color display. The predictions with regard to spatial proximity were less clear. In the previous two experiments, the dominant feature affecting performance had been the proximity between relevant and irrelevant channels. In the absence of the latter, our interest was in whether the principle held in a clutter-free display.

Method

Design and Procedure

The design and procedure for Experiment 3 were identical to those of Experiments 1 and 2 with two exceptions: (1) the clutter was removed, leaving a display of only the three relevant indicators, and (2) a common color display (used in Experiment 2) that coded all three relevant variables in the same color (magenta) was contrasted with a separate color display (used in Experiment 1) in which the airspeed indicator was white while the bank and flaps indicators were magenta and blue, respectively. As with the previous two experiments, spatial proximity was varied among three levels: close, intermediate, or distant.

Subjects

Ten subjects (undergraduates attending the University of Illinois) were paid to participate in the experiment. All subjects had normal or corrected-to-normal vision and were screened for their ability to perceive the actual color used in the experiment (by self report). The subjects had no flight experience.

Results

Color Proximity

Figure 9 plots the error data for the two types of tasks (integration and focused attention) as a function of the color coding schemes. The data are

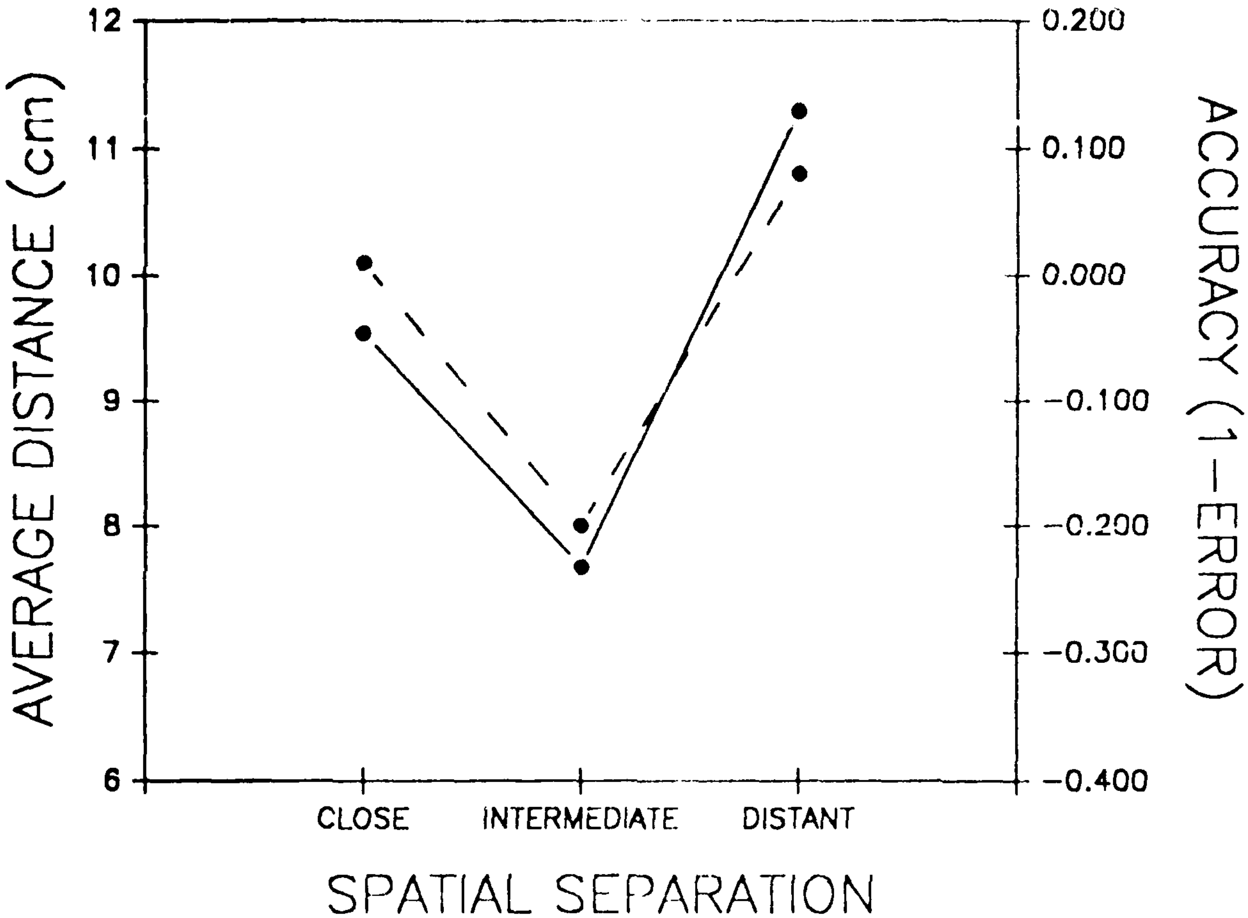

Figure 8. Mean distance between relevant and irrelevant indicators at each spatial separation level (solid line). (The dashed line plots accuracy [1-error], reflecting the similarity between the two functions.)

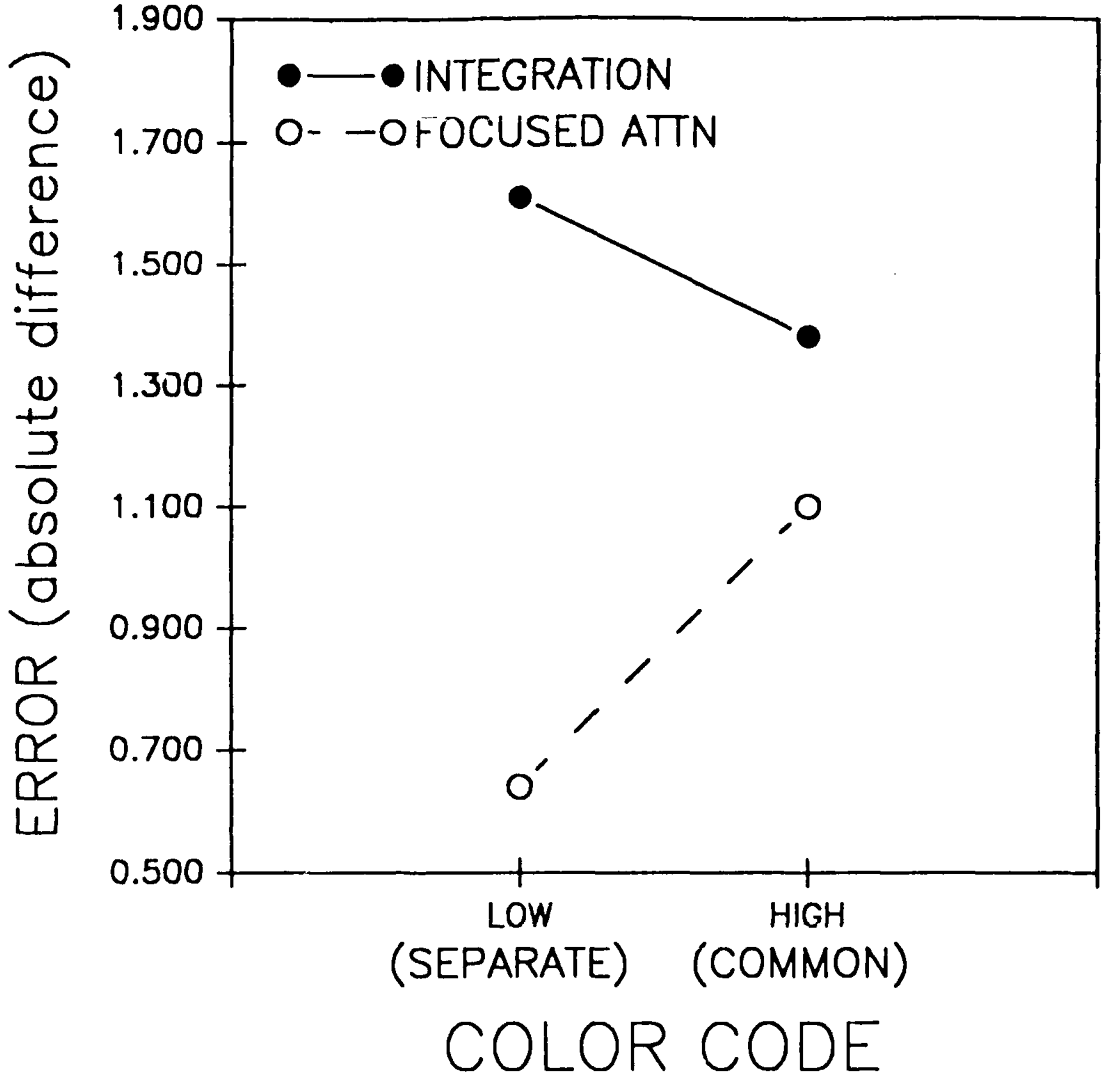

Figure 9. Effect of color code and task type on error (Experiment 3).

averaged across levels of spatial separation, while the focused attention data are averaged across indicator type. There was no main effect of display color code ($\underline{F}[1,9] = 3.45$, $\underline{p}<.10$), but a significant Display x Task interaction supports the compatibility proximity hypothesis, $\underline{F}(3,27) = 7.15$, $\underline{p}<.01$. Although the trend is in the predicted direction, pairwise contrasts within this interaction revealed that integration performance was not significantly affected by the common color display ($F[1,9] = 4.07$, $p=.08$); while focused attention performance was facilitated by the separate color display ($F[1,9] = 5.84$, $p<.05$). In addition, a significant Color x Indicator x Recall interaction occurred in the error data. This interaction revealed that the magnitude of the effects of the separate color display on focused attention performance was different for the three indicators, $F(2,18) = 6.48$, $p<.01$. Although recall of each indicator was facilitated by the separate color scheme, this benefit (in error reduction) was greatest when attention had to be focused on the flaps indicator, the variable that was given the least weight in the stall equation. Analysis of the reaction time data failed to reveal any significant effects between the two color formats.

Spatial Proximity

The manipulation of spatial proximity did not reveal significant accuracy effects for either integration ($\underline{F}[2,18] = .59$, ns) or focused attention performance ($\underline{F}[2,18] = 1.86$, $\underline{p}>.18$), although the data showed a trend toward a pattern predicted by the compatibility of proximity principle. The most important feature of the current data was the elimination of the disruptive effect of the intermediate spatial separation condition, which had been found in Experiments 1 and 2.

Reaction time data failed to reveal significant effects for manipulations of spatial proximity for either focused attention or integration performance.

For focused attention performance, a significant Color x Space interaction for the accuracy data occurred showing that the benefits of the separate color display on focused attention performance changed as a function of the spatial distance between indicators, $\underline{F}(2,18) = 5.02$, $\underline{p}<.05$. The magnitude of this benefit (in error reduction) was greatest when the indicators were displayed in close spatial proximity. Thus, the separate color code helped focused attention performance most in the close spatial proximity condition where overall focused attention performance was poorest.

Discussion

The results of Experiment 3 again revealed that manipulations of color proximity produced an interaction consistent with the compatibility of proximity principle. Thus, integration task performance was best with the common color display while focused attention performance was best with the separate color display, reinforcing the conclusions offered by the first two experiments.

Results from manipulations of spatial proximity suggested trends in the direction predicted by the proximity compatibility principle and provided an important contrast to the previous two experiments which incorporated clutter. That is, the disrupting effect in the intermediate separation display produced when that clutter was closest to the relevant indicators was eliminated. Likewise, the facilitating effect in the distant separation display when clutter was farthest from the relevant indicators was also eliminated.

It is important to note that the reaction time results provided the significant Display x Task interactions in Experiment 1, while in Experiments 2 and 3 the significant Display x Task interactions were provided by accuracy. In both experiments, however, the dependent measure not showing the Display x Task interaction either showed a consistent trend at a nonsignificant level or did not discriminate between conditions. In other words, the pattern of results was not produced by a speed-accuracy tradeoff.

GENERAL DISCUSSION

The results from this study have contributed to our understanding of the perception of relations among display attributes and its effect on the corresponding mental processing of the information these attributes convey. More specifically, the three experiments presented here have suggested some important conclusions bearing on the effects of color and space as they pertain to the proximity compatibility principle.

Color Proximity

Collectively, the results from the three experiments replicate previous findings that color strongly affects the perceptual organization of display elements. The present results suggest that this organization affects performance in a way consistent with the proximity compatibility principle. That is, the use of separate colors facilitates the focused attention recall of those variables that are uniquely colored in the display, and concurrently degrades the ability to integrate information. Further, the use of a common color for all relevant indicators facilitates their integration, but at the apparent cost of recalling their unique identities. Hence, color is helpful if it is logically related to the operator's task (Hale & Billmayer, 1988).

Spatial Proximity

The current results concerning spatial separation are consistent with the vast amount of literature that suggests that physical space is the predominant factor in the perceived organization of an information display (e.g., Holahan et al., 1978; Tullis, 1983). However, within the context of a cluttered display, spatial separation per se did not conform to the proximity compatibility principle. Instead, the effect of spatial separation was manifest as measured by the distance between relevant and irrelevant information and was not task dependent. Thus, spatial separation and organization appears to be most effective at an early processing stage in segregating "signals" from "noise" throughout the display. When this segregation is adequately achieved (as in the distant spatial configuration in Experiments 1 and 2), performance on both focused attention and integration tasks will benefit. However, when the proximity between relevant and irrelevant information makes this segregation difficult (as in the intermediate spatial configuration in Experiments 1 and 2), the ability to integrate information and focus attention on particular variables is impaired. The elimination of these effects in Experiment 3, when no clutter was present, directly implicates the role of clutter proximity as a disruptive feature in multi-element display processing.

Implications for Display Design

The applied implications of these results should be highlighted in terms of two emerging trends in multi-element display design. First, the rapid emergence of color as a feature in electronic displays is probably an inevitable trend because of its inherent attractiveness, in spite of the fact that few objective advantages have been demonstrated by empirical work in task environments more complex than visual search (Hale & Billmayer, 1988). Hence, while color has been shown to enhance performance in a wide range of tasks such as detection, discrimination, identification, and classification (Berggrund et al., 1984), its efficacy has not been established regarding performance in more complex, cognitive activities. The ambivalence of evidence results in part from an absence of theory-based principles (derived from information processing-based task analyses) used to identify those ways in which color coding can facilitate multi-element display processing. Clearly, the current results have helped to validate the utility of such principles.

Secondly, the rate and complexity of data presentation in current high-demand environments "is such that complex formats are inevitable, with display clutter becoming a major problem" (Martin, 1984, p. 7.1). The data from the present experiments provide qualitative guidelines concerning the role of display clutter in information extraction. The data go further in suggesting, as shown in Figure 8, the foundation for quantitative models of the effects of display clutter on performance. Such models have been proposed by Tullis (1983) and Palmiter and Elkerton (1987) to offer important benefits to the early stages of display design, and are needed in emerging model-based tools for aviation system design (Elkind, Card, Hochberg, & Huey, 1989).

Guidelines for Future Research

The approach advocated in this experiment serves as a guideline for future research concerning display processing and design. Paramount to good display theory is its validation as an "evaluation tool" that serves to predict the advantages and disadvantages of introducing a new display format without the costs of actual implementation (Tullis, 1983). Unfortunately, little evidence exists to directly validate the proximity compatibility principle regarding real-world operational tasks. Much of the previous research has approached the validation of this principle by focusing on a display concept (e.g., space) and then fabricating a task to test the principle. The task, however, is often somewhat arbitrary and maps only vaguely into real-world operational settings. This approach is well chosen to serve good theory, but may not provide generalizable principles, since the tasks may be specifically selected to capitalize on the display dimensions chosen. In contrast to this approach, the display designer outside of the laboratory must work in the opposite sequence. The designer is given a task first, with a fairly restricted set of constraints; and then is asked to develop a display design that best supports focused attention, independent processing, or information integration, as these demands are revealed by the appropriate task analysis.

In the current experiments, the approach described is intermediate between these two approaches. That is, we have chosen a real-world task, and then fabricated its components to more easily examine the principle in a laboratory setting. However, the task chosen is quite prototypical of many information processing tasks in the cockpit setting and in other complex systems, and thus illustrates how the system designer might approach the display design problem. Clearly, the tasks performed in laboratory settings should be heavily guided by the kinds of tasks performed in real-world operational settings. This approach first entails an analysis of the information processing required by the system operators. Following that, one must examine the display formats that best support these information processing requirements. This is the approach that we have taken here and advocate for future research in this area.

CONCLUSION

In conclusion, two objectives were served by the current study: (1) to test the proximity of compatibility principle when applied to dimensions of space and color, and (2) to demonstrate the applicability of the proximity compatibility principle to real-world issues of display design. In meeting the first of these objectives, the results demonstrated both the breadth of and the constraints on the principle. When proximity was defined by spectral color, the principle was clearly upheld regarding both its implications for focused attention and information integration. At the same time, constraints on the breadth of the principle were identified with regard to spatial proximity. Within the range of distances used here, and in the presence of irrelevant visual information, spatial closeness neither facilitated integration nor hindered attention focusing. Rather, it appears as an organizing factor that can segregate relevant from nonrelevant sources of information. Furthermore, the current results appear to support the development of quantitative guidelines concerning the role of clutter in visual display processing; a notion that has been voiced by other researchers (Palmiter & Elkerton, 1987; Tullis, 1983).

The second objective was met by demonstrating the applicability of the principle, as defined by color, to a valid aviation problem: How to convey to the pilot, in a rapid glance under time pressure, information regarding the degree of stability of the aircraft under his control. The current research shows directly how this principle, relating the physical proximity of display elements to the mental proximity of the task information processing requirements, has important implications for the design of multi-element displays.

REFERENCES

Aretz, A. J., & Calhoun, G. L. (1982). Computer generated pictorial stores management displays for fighter aircraft. In Proceedings of the Human Factors Society 26th Annual Meeting (pp. 455-459). Santa Monica, CA: The Human Factors Society.

Banbury, J. R. (1984). Electronic displays for use in military cockpits. In C. P. Gibson (Ed.), Proceedings of the Workshop on Colour Coded vs Monochromatic Electronic Displays (pp. 1.1-1.19). Farnborough, England: Royal Aircraft Establishment.

Barnett, B., & Wickens, C. D. (1988). Display proximity in multicue information integration: The benefits of boxes. Human Factors, 29(5), 170-195.

Berggrund, U., Derefeldt, G., Hedin, C. E., & Marmolin, H. (1984). Colour coded vs monochrome situational maps. In C. P. Gibson (Ed.), Proceedings of the Workshop on Colour Coded vs Monochromatic Electronic Displays (pp. 11.1-11.10). Farnborough, England: Royal Aircraft Establishment.

Boff, K. R., & Lincoln, J. (1988). Engineering data compendium: Human perception and performance. Wright-Patterson AFB, OH: AAMRL.

Boles, D. B., & Wickens, C. D. (1987). A comparison of homogeneous and heterogeneous display formats in information integration and nonintegration tasks (Technical Report EPL-83-6/ONR-83-6). Urbana, IL: University of Illinois, Engineering-Psychology Research Laboratory.

Brindle, J. H., & Mulley, W. G. (1984). The pursuit of color from a display technology perspective. In C. P. Gibson (Ed.), Proceedings of the Workshop on Colour Coded vs Monochromatic Electronic Displays (pp. 4.1-4.17). Farnborough, England: Royal Aircraft Establishment.

Broadbent, D. E. (1982). Task combination and selective intake of information. Acta Psychologica, 50, 253-290.

Buffett, A. R. (1986). Visual perception of luminance contrast: Implications for the assessment of subjective image quality. Displays, 7(1), 30-36.

Bundesen, C., & Pedersen, L. F. (1983). Color segregation and visual search. Perception & Psychophysics, 33(5), 487-493.

Carswell, C. M., & Wickens, C. D. (1987). Information integration and the object display: An interaction of task demands and display superiority. Ergonomics, 30, 511-528.

Carswell, C. M., & Wickens, C. D. (1988). Comparative graphics: History and applications of perceptual integrality theory and the proximity compatibility hypothesis (Technical Memorandum 8-88). Aberdeen Proving Ground, MD: U.S. Army Human Engineering Laboratory.

Carswell, C. M., & Wickens, C. D. (1988). Technical report ARL-88-2/AHEL-88-1. Urbana, IL: University of Illinois.

Carter, R. C. (1982). Search time with a color display: Analysis of distribution functions. Human Factors, 24(2), 203-212.

Christ, R. E. (1975). Review and analysis of color coding research for visual displays. Human Factors, 17(6), 542-570.

Cleveland, W. S. (1985). The elements of graphing data. Monterey, CA: Wadsworth.

Cleveland, W. S., & McGill, R. (1984). Graphical perception: Theory, experimentation, and application to the development of graphic methods. Journal of the American Statistical Association, 70, 531-554.

Curry, R. E. (1985). The introduction of new cockpit technology: A human factors study (NASA Technical Memorandum 86659). Moffett Field, CA: NASA-Ames Research Center.

Elkind, J. I., Card, S. K., Hochberg, J., & Huey, B. M. (Eds.). (1989). Human performance models for computer-aided engineering. Washington, DC: National Academy Press.

Eriksen, B. A., & Eriksen, C. W. (1974). Effects of noise letters upon the identification of a target letter in a nonsearch task. Perception & Psychophysics, 16, 143-149.

Eriksen, C. W., & Hoffman, J. E. (1972). Some characteristics of selective attention in visual perception determined by vocal reaction time. Perception & Psychophysics, 11, 169-171.

Eriksen, C. W., & Hoffman, J. E. (1973). The extent of processing of noise elements during selective encoding from visual displays. Perception & Psychophysics, 14, 155-160.

Eriksen, C. W., & St. James, J. D. (1986). Visual attention within and around the field of focal attention: A zoom lens model. Perception & Psychophysics, 40, 225-240.

Eriksen, C. W., & Yeh, Y. Y. (1985). Allocation of attention in the visual field. Journal of Experimental Psychology: Human Perception and Performance, 11, 583-597.

Godfrey, G. W. (1982). Principles of display illumination techniques for aerospace vehicle crew stations (2nd ed.). Tampa FL: Aerospace Lighting Institute.

Goettl, B. P., Kramer, A. F., & Wickens, C. D. (1986). Display format and the perception of numerical data. In Proceedings of the Human Factors Society 30th Annual Meeting (pp. 450-454). Santa Monica, CA: The Human Factors Society.

Hale, S. L., & Billmayer, H. J. (1988). Use of color CRTs in aircraft cockpits: A literature search. (Technical Note 3-88). Aberdeen Proving Ground, MD: U.S. Army Human Engineering Laboratory.

Harwood, K., Wickens, C. D., Kramer, A. F., Clay, D., & Liu, Y. (1986). Effects of display proximity and memory demands on the understanding of dynamic multidimensional information. In Proceedings of the Human Factors Society 32nd Annual Meeting (pp. 786-789). Santa Monica, CA: The Human Factors Society.

Holahan, C. J., Culler, R. E., & Wilcox, B. L. (1978). Effects of visual distraction on reaction time in a simulated traffic environment. Human Factors, 20(4), 409-413.

Hudson, P. T. W. (1984). Encoding information in displays: Colour versus non-coloured methods and their uses, or, what you can do extra with a colour display. In C. P. Gibson (Ed.), Proceedings of the Workshop on Colour Coded vs Monochromatic Electronic Displays (pp. 34.1-34.13). Farnborough, England: Royal Aircraft Establishment.

Humphreys, G. W. (1981). On varying the span of visual attention: Evidence for two modes of spatial attention. Quarterly Journal of Experimental Psychology, 33A, 17-31.

Huntoon, R. B. (1985). Displays, deja vu. Aviation, Space, and Environmental Medicine, 56(2), 131-137.

Kramer, A. F., Wickens, C. D., Goettl, B., & Harwood, K. (1986). Extraction and comprehension of data relations: An application of the display proximity principle. In R. Ginnett (Ed.), Proceedings of the 10th Symposium on Psychology in the Department of Defense (DOD). Colorado Springs, CO: United States Air Force Academy.

Laberge, D. (1983). Spatial extent of attention to letters and words. Journal of Experimental Psychology: Human Perception and Performance, 9, 371-379.

Laycock, J., & Chorley, R. A. (1980). The electro-optical display/visual system interface: Human factors considerations. Agardograph, 255, 311-315.

Lovesey, E. J. (1986). The use of an air combat simulator in cockpit display development. In Proceedings of a Symposium entitled "Developments in Information Presentation to Aircrew", Vol. 2 (pp. 13.1-13.9). North Luffenham, England: Air Standardization Coordinating Committee.

Lyons, J. W., & Roe, G. (1980). The influence of visual requirements on the design of military cockpits. Agardograph, 255, 121-129.

Mahaffey, D. L., Horst, R. L., & Munson, R. C. (1986). Behavioral comparison of the efficacy of bar graphs and polar graphics for displays of system status. In IEEE Proceedings of the International Conference on Systems, Man, and Cybernetics, 1514-1519.

Martin, K. W. (1984). Display visual performance prediction in a military cockpit environment. In C. P. Gibson (Ed.), Proceedings of the Workshop on Colour Coded vs Monochromatic Electronic Displays (pp. 7.1-7.15). Farnborough, England: Royal Aircraft Establishment.

Moar, I. (1977). Spatial organization in the free recall of pictorial stimuli. _Quarterly Journal of Experimental Psychology_, _29_(4), 699-708.

Moray, N. (1988). Mental workload since 1979. _International Review of Ergonomics_, _2_, 123-150.

Narborough-Hall, C. S. (1985). Recommendations for applying colour coding to air traffic control displays. _Displays_, _6_(3), 131-137.

Palmiter, S. L., & Elkerton J. (1987). Evaluation metrics and a tool for control panel design. In _Proceedings of the Human Factors Society 33rd Annual Meeting_ (pp. 1123-1127). Santa Monica, CA: The Human Factors Society.

Pamperin, K. L., & Wickens, C. D. (1987). The effects of modality and stress across task type on human performance. In _Proceedings of the Human Factors Society 33rd Annual Meeting_ (pp. 514-518). Santa Monica, CA: The Human Factors Society.

Peterson, R. J., Banks, W. W., & Gertman, D. J. (1981). Performance-based evaluation of graphic displays for nuclear power plant control rooms. In _Proceedings of the Conference on Human Factors in Computing Systems_ (pp. 182-189). Gaithersburg, MD: ACM.

Podgorny, P., & Shepard, R. N. (1983). Distribution of visual attention over space. _Journal of Experimental Psychology: Human Perception and Performance_, _9_(3), 380-393.

Polson, M., Wickens, C. D., Klapp, S. T., & Colle, H. (1989). Human interactive information processes. In P. Hancock and M. H. Chignell (Eds.), _Intelligent interfaces: Theory, research, and design_ (pp. 129-164). Elsevier Science.

Rasmussen, J. (1986). _Information processing and human-machine interaction: An approach to cognitive engineering_. New York: North-Holland.

Reising, J. M., & Calhoun, G. L. (1982). Color display formats in the cockpit: Who needs them? In _Proceedings of the Human Factors Society 26th Annual Meeting_ (pp. 446-449). Santa Monica, CA: The Human Factors Society.

Reising, J. M., Emerson, T. J., & Aretz, A. J. (1984). Computer-generated formats for advanced fighter cockpits. In C. P. Gibson (Ed.), _Proceedings of the Workshop on Colour Coded vs Monochromatic Electronic Displays_ (pp. 31.1-31.9). Farnborough, England: Royal Aircraft Establishment.

Santucci, G., Menu, J. P., & Amalberti, R. (1984). Single or multicolor displays: How to choose. In C. P. Gibson (Ed.), _Proceedings of the Workshop on Colour Coded vs Monochromatic Electronic Displays_ (pp. 24.1-24.13). Farnborough, England: Royal Aircraft Establishment.

Silverstein, L. D. (1987). Human factors for color display systems: Concepts methods, and research. _Color and the computer_. Academic Press.

Snyder, H. L. (1984). Effect of color contrast on legibility of airborne displays. In C. P. Gibson (Ed.), Proceedings of the Workshop on Colour Coded vs Monochromatic Electronic Displays (pp. 27.1-27.19). Farnborough, England: Royal Aircraft Establishment.

Statler, I. C. (1984). Military pilot ergonomics. In Proceedings of the AGARD Aerospace Medical Panel Symposium on Human Factors Considerations in High Performance Aircraft (pp. 13-23). Williamsburg, VA: North Atlantic Treaty Organization.

Stokes, A., & Wickens, C. D. (1988). Aviation displays. In E. Wiener and D. Nagel (Eds.), Human Factors in Aviation (pp. 387-422). New York: John Wiley.

Taylor, R. M. (1985). Colour design in aviation cartography. Displays, 6(4), 187-201.

Tullis, T. S. (1983). The formatting of alphanumeric displays: A review and analysis. Human Factors, 25(6), 657-682.

Wertheimer, M. (1958). Principles of perceptual organization. In D. C. Beardslee & M. Wertheimer (Eds.), Readings in Perception. Princeton, NJ: Van Nostrand.

Wickens, C. D. (1984). Engineering psychology and human performance. Glenview, IL: Scott, Forsemen Publishing Company.

Wickens, C. D. (1987). Ergonomic-driven displays: A set of principles of the perceptual-cognitive interface. In Proceedings of the 4th Mid-Central Ergonomics/Human Factors Conference.

Wickens, C. D. (1988). Human interface design. In Seminar Lecture Notes of the Society for Information Display Annual Meeting, Vol. 1 (pp. 7.1-7.31).

Wickens, C. D., & Flach, J. M. (1988). Information processing. In E. Wiener and D. Nagel (Eds.), Human Factors in Aviation (pp. 11-149). New York: John Wiley.

Wiener, E. L. (1985). Human factors of cockpit automation: A field study of flight crew transition (NASA Contractor Report CR-177333). Moffett Field, CA. NASA-Ames Research Center.

Wiener, E. L., & Curry, R. E. (1980). Flight-deck automation: Promises and problems. Ergonomics, 23, 995-1011.

Wierwille, W. W., Rahimi, M., & Casali, J. G. (1985). Evaluation of 16 measures of mental workload using a simulated flight task emphasizing mediational activity. Human Factors, 27(5), 489-502.

Woods, D. D., & Roth, E. M. (1988). Cognitive engineering: Human problem solving with tools. Human Factors, 30(4), 415-430.